ELSEVIER

Theoretical Computer Science 164 (1996) 223–252

Theoretical Computer Science

An application of Hajós factorizations to variable-length codes

Clelia De Felice [1]

Dipartimento di Informatica ed Applicazioni, Università di Salerno, 84081 Baronissi (SA), Italy

Received September 94; revised October 95
Communicated by D. Perrin

Abstract

We give a new characterization of some factorizations of finite cyclic groups described by Hajós. As a consequence, we prove that the embedding problem for a particular class of codes is decidable. In particular, we link this problem for codes $C \subseteq a^*b \cup ba^*$ to Schützenberger's factorization conjecture, via Hajós factorizations.

1. Introduction

A *variable-length code* is the base of a free submonoid of A^*, that is a set of words C, over an alphabet A, such that any message written by using the words in C as code words can be uniquely decoded. The algebraic theory of codes was initiated by Schützenberger and then extensively developed by him and his school [2, 41, 42].

An important property of a code is maximality, a *maximal* code being a code which is not properly included in any other code over the same alphabet. By Zorn's lemma, any code C can be embedded in a maximal one (a *completion* of C). But this result is no longer true if we restrict ourselves to finite codes: there exist finite codes having no finite completion [3, 33]. The smallest known example is the code $\{a^5, ba^2, ab, b\}$ [2, 33].

Therefore, an intriguing problem in the theory of variable-length codes is to characterize those finite codes that can be embedded in a finite maximal one. Actually, it is not even known whether this property is effectively decidable. This problem is a particular case of a more general one [9]: given a code C which has some property P, does a maximal code C' exist which contains C and retains the same property P? Despite its difficulty, there are interesting results for some particular classes of codes, namely for *recognizable* codes [20], *prefix* or *suffix* codes [2], *biprefix* codes [30, 44], codes

[1] Partially supported by ESPRIT-BRA Working Group 6317 *ASMICS*, Project 40% MURST "*Efficienza di Algoritmi e Progetto di Strutture Informative*" and Project 60% MURST.

SSDI 0304-3975(95)00223-5

having *bounded deciphering delay* [13]. Moreover, another direction of research has been started which considers automata as tools for solving this problem of completion of codes [1, 10, 28].

In this paper, we will only consider the embedding problem for finite codes and we will restrict ourselves to a two-letter alphabet $A = \{a, b\}$. The partial known results about this problem are some methods for constructing families of codes having no finite completion [17, 26, 27, 33, 34]. In particular it is possible to get a finite code C containing a^n and having no finite completion for any $n \geqslant 2$ [26, 27]. But to decide the existence of a finite completion is a difficult problem even when the cardinality of the code is very small (see [19, 34]). No algorithm exists for this decision problem and, consequently, no procedure exists for embedding a finite code in a maximal one. All known codes having no finite completion are constructed thanks to a relation between maximal codes and factorizations of cyclic groups which has been proved in [34].

This paper is a survey of these two topics, embedding of codes and factorizations, with some new results. Precisely, we will present some cases in which the existence of a finite completion is decidable and, in the case of a positive answer, a finite completion is also effectively constructed (Section 7). Our results are based on the previous theorem in [34] and also on a new characterization of a class of factorizations of $\boldsymbol{Z}_n$, discovered by Hajós (Section 5).

We recall that a pair (T, R) of subsets of $\boldsymbol{N}$ is a *factorization* of $\boldsymbol{Z}_n$ if for any $z \in \{0, \ldots, n-1\}$ there exists a unique pair (t, r), with $t \in T$ and $r \in R$, such that $t + r = z \pmod{n}$.

The general structure of the factorizations of $\boldsymbol{Z}_n$ is unknown. By solving a conjecture proposed by Minkowski, Hajós conjectured that any factorization (T, R) of a finite abelian group G was periodic (i.e. there exists $g \in G \setminus 0$ such that $g + T = T$). This conjecture was false. Subsequently, cyclic groups were classed in *good* groups (verifying Hajós conjecture) and *bad* groups (the others) [21].

In [23], Hajós gives a method for constructing a class of periodic factorizations containing all factorizations of a good group. Theorem 3.2 gives a new characterization of them. A by-product of this theorem is their simple recursive construction which gives a complete description of the structure of these factorizations (Proposition 5.1). The new characterization is related to the construction of some codes verifying the so-called Schützenberger's *factorization conjecture.*

Using our construction of Hajós factorizations we can prove the other results, concerning partial answers to the embedding problem, for codes with the form $C = a^n \cup a^P b \cup b a^Q$, $P, Q \subset \boldsymbol{N}, b \in C$. They stress different aspects of this problem for finite codes.

First, we prove that one can effectively construct a finite completion for codes C with a pair (P, Q) which is embeddable in a Hajós factorization of $\boldsymbol{Z}_n$ (see Proposition 7.2). If $\boldsymbol{Z}_n$ is a good group, then C has a finite completion if, and only if, (P, Q) has the previous property (see Corollary 7.1). For general groups $\boldsymbol{Z}_n$, this condition on (P, Q) is necessary and sufficient for the existence of a particular embedding of C. Namely,

a finite completion which verifies the factorization conjecture (Proposition 7.3). In this way we link these two problems about codes. (See Sections 6 and 7 for some more details.)

Note that the existence of $n \in \mathbf{N}$, such that (P,Q) can be embedded in a Hajós factorization of $\mathbf{Z}_n$, is a decidable property for a pair (P,Q). So, Propositions 7.2 and 7.3 can be reformulated for codes with the form $C = a^P b \cup ba^Q$ and $b \in C$.

This paper is organized as follows. Section 2 contains some preliminaries. In Section 3 we recall the definitions and known results of factorizations of a cyclic group which will be subsequently referred to. The same is done for codes in Section 4. We give our construction of Hajós factorizations in Section 5. In Section 6 we survey some conjectures, two concerning the general structure of the factorizations of cyclic groups, one concerning codes. Then, we relate them to each other. In Section 7, we prove partial results about the embedding problem for finite codes. Finally, in the Appendix we gather some technical results. An extended abstract of all these results, without proofs, has already been communicated [16].

2. Preliminaries

In the next section we recall the definitions and previous results about the factorizations of a cyclic group (Section 3). Subsequently, we will introduce the basic properties of codes which we require (Section 4). For both these topics we need the techniques and notations of the polynomials and the formal power series in non-commutative variables [4].

Let A be a finite alphabet and A^* be the *free monoid* generated by A. 1 is the empty word and A^+ is the set $A^* \setminus 1$. We denote $K\langle A\rangle$ the semiring of the *non-commutative polynomials* generated by A over a semiring K. Any finite subset X of A^* will be identified with its *characteristic polynomial*: $\underline{X} = \sum_{x \in X} x$. For a polynomial $P \in \mathbf{Z}\langle A\rangle$, $P \geqslant 0$ means that P has non-negative coefficients. For any word $w \in A^*$, (P,w) denotes the coefficient of w in P.

Let M be a finite multiset of $\mathbf{N}$. We denote a^M the polynomial $\sum_{n \in \mathbf{N}} (M,n)a^n \in \mathbf{N}\langle a\rangle$. The computation rules are

$$a^{\emptyset} = 0, \quad a^0 = 1, \quad a^{M \cup N} = a^M + a^N, \quad a^{M+N} = a^M a^N.$$

The symbols $\cup$ and $+$ mean union and addition for multisets.

Recall that a word $x \in A^*$ is a *prefix* (resp. *proper prefix*) of a word $w \in A^*$ if $w = xy$, with $y \in A^*$ (resp. $y \in A^+$).

3. Factorizations of cyclic groups

In this section, we will recall the definitions and known results about the factorizations of abelian groups. We will also announce one of the main results of this

paper (see Theorem 3.2). We consider first some general results about factorizations (Section 3.1). Subsequently, we consider periodic factorizations (Section 3.2) and we end the section with some results about Hajós factorizations (Section 3.3).

3.1. Definitions

In the following, group is used to mean a finite abelian group, which we denote by G. Moreover, we suppose that the law of composition is written additively. The notion of factorization of a group was introduced for the first time by Hajós when he solved one of Minkowski's conjectures by giving it a group-theoretical formulation [21, 23, 24]. We recall that, given a finite abelian group G, a sequence $S_1, \ldots, S_k$ of subsets of G is a *factorization* of G (or G is the *direct sum* of its subsets S_i) if each element of G may be written uniquely as a sum with just one term from each S_i. Then S_i is called a *factor* of G.

For our aims, we deal mainly with factorizations of cyclic groups. As usual, we realize the cyclic group of order n as the factor group $\boldsymbol{Z}_n$ of the integers modulo n. For the relation with codes, we will always consider *positive* representatives of its classes. So, when it is possible, we consider subsets of $\boldsymbol{N}$ and relations in $\boldsymbol{N}$, instead of the corresponding situations in $\boldsymbol{Z}_n$.

Moreover, in this paper we will focus our attention on the factorizations of $\boldsymbol{Z}_n$ by two factors, defined as follows.

Definition 3.1. A pair (T,R) of subsets of $\boldsymbol{N}$ is a factorization of $\boldsymbol{Z}_n$ if for any $z \in \{0,\ldots,n-1\}$ there exists a unique pair (t,r), with $t \in T$ and $r \in R$, such that $t + r = z \pmod{n}$.

In this case we can define the set H of the *holes* of (T,R) as follows:

$$H = \{q \in \boldsymbol{N} \mid \forall q' \in T + R \quad q \equiv q' \pmod{n} \Rightarrow q' > q\}.$$

So, if $n = p$ is a prime number, a pair (T,R) of subsets of $\boldsymbol{N}$ is a factorization of $\boldsymbol{Z}_p$, if and only if we have $T = \{t\}$ and $R = \{0,\ldots,p-1\} \pmod{p}$. As an example, $(\{2\},\{3,7,8\})$ is a factorization of $\boldsymbol{Z}_3$. In this case, we have $H = \{0,1,2,3,4,6,7\}$.

For our aims, it is also useful to consider an equivalent definition of factorization of cyclic groups in terms of polynomials in $\boldsymbol{N}\langle a\rangle$. We get this using the following proposition.

Proposition 3.1. *A pair (T,R) of subsets of $\boldsymbol{N}$ is a factorization of $\boldsymbol{Z}_n$ if and only if there exists a finite subset H of $\boldsymbol{N}$ such that*

$$a^T a^R = \left(\frac{a^n - 1}{a - 1}\right)(1 + a^H(a-1)). \tag{3.1}$$

H is the set of the holes of (T,R).

Proof. If (T,R) verifies (3.1), then we have $a^T a^R \equiv 1+a+\ldots+a^{n-1} \pmod{(a^n-1)}$. It is classically known that this relation implies that (T,R) is a factorization of $\boldsymbol{Z}_n$ ([7, 23], see also Lemma 3.2(ii) in [15]).

Conversely, suppose that (T,R) is a factorization of $\boldsymbol{Z}_n$. So, for any $z \in \{0,\ldots,n-1\}$ there exists a unique pair (t,r), with $t \in T$ and $r \in R$, such that $t+r = z+s_z n$, where $s_z \in \boldsymbol{N}$. Then, the set H of the holes of (T,R) is given by

$$H = \{z+sn \mid z \in \{0,\ldots,n-1\}; \quad s \in \{0,\ldots,s_z-1\}\}.$$

Using this relation, (3.1) easily follows. □

The structure of these pairs (T,R) is still unknown except in some special cases.

For instance in case $H = \emptyset$ these pairs are described by Krasner and Ranulac [25]. They proved that the polynomial $1+a+\ldots+a^{n-1}$ is a product of two polynomials with real non-negative coefficients if and only if we have

$$\frac{a^n-1}{a-1} = a^I a^J,$$

where I, J are subsets of $\boldsymbol{N}$. We call such a pair (I,J) a *Krasner factorization*. In the same paper [25], the authors characterized the structure of these factorizations. Any pair (I,J) can be obtained by taking a chain of divisors of n

$$k_0 = 1 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n$$

by writing the equation

$$\frac{a^n-1}{a-1} = \frac{a^n-1}{a^{k_{r-1}}-1} \cdots \frac{a^{k_2}-1}{a^{k_1}-1} \frac{a^{k_1}-1}{a-1}$$

and by setting

$$\forall q \in \{1,\ldots,r\} \quad P_q = \frac{a^{k_q}-1}{a^{k_{q-1}}-1}, \quad a^I = \prod_{q \text{ even}} P_q, \quad a^J = \prod_{q \text{ odd}} P_q.$$

In [8] it is proven that this proposition is again true for *any finite sum* of multisets $I+J+\ldots+K = \{0,\ldots,n-1\}$. Moreover a recursive version of this algorithm can be found in [14, 26]. More precisely (I,J) is a Krasner factorization of $\boldsymbol{Z}_n$ if and only if there exists a divisor k of n and a Krasner factorization (I_1,J_1) of $\boldsymbol{Z}_{n/k}$ such that

$$I = kJ_1, \quad J = kI_1 + \{0,\ldots,k-1\}.$$

(As a matter of fact k equals the smallest divisor k_1, greater than 1, in the chain of divisors of n.) For instance, $(\{0\},\{0,1\})$ is a Krasner factorization of $\boldsymbol{Z}_2$.

3.2. Periodic factorizations

A class of factorizations larger than Krasner's family is obtained by considering the so-called periodic factorizations.

Denote $\langle g\rangle$ the subgroup of G generated by $g \in G$. We recall that a subset S of G is *periodic* if there exists g in $G \setminus 0$ such that $g + S = S$. In other words, S is periodic if and only if it has the form $S = \langle g\rangle + S'$, for some non-zero element g and for some subset S' of G [21].

A factorization (T,R) of G is *periodic* if at least one factor is periodic. Hajós conjectured that any factorization of a finite abelian group was periodic. This conjecture was false. Thereafter, cyclic groups were classed in good groups and bad groups. A group is *good* if *any* of its factorization is periodic, otherwise it is *bad*. For instance $\boldsymbol{Z}_p$ is a good group.

When we consider the particular case of the cyclic groups, one has the following obvious characterization of the periodic factorizations, which we need in the sequel.

Lemma 3.1 [21]. *(T,R) is a periodic factorization of $\boldsymbol{Z}_n$ if and only if there exists a divisor q of n, $q \neq n$, and a factorization (T,S) of $\boldsymbol{Z}_q$ such that R is the direct sum of S and $\{0,q,\ldots,(n/q-1)q\}$.*

In a sequence of different papers, good cyclic groups were characterized [6, 7, 21, 23, 24, 37–39]. They are $\boldsymbol{Z}_{p^\alpha q}$, $\boldsymbol{Z}_{p^2q^2}$, $\boldsymbol{Z}_{p^2qr}$, $\boldsymbol{Z}_{pqrs}$ and their subgroups, where p, q, r, s are different primes.

Moreover, the simplest non-periodic factorization known at present is the factorization (T,R) of $\boldsymbol{Z}_{72}$ given by

$$T = \{0,8,16,18,26,34\}, \quad R = \{0,6,12,36,42,48,1,5,25,29,49,53\}$$

(see [7]).

We end this section by stressing that Krasner factorizations (I,J) are periodic. Indeed, I (or J) can be written as $I = I' + \{0,k_{r-1},\ldots,k_{r-1}(n/k_{r-1}-1)\}$, where k_{r-1} is a divisor of n.

3.3. *Hajós factorizations*

An interesting class of factorizations was found by Hajós. In [23], Hajós gave a method which gives a class of periodic factorizations which he claimed contains all factorizations of a good group. However, in [37] Sands pointed out that this method needed a slight correction. We report this corrected version below.

Let $S = \{s_1,\ldots,s_q\}$, T be subsets of G. We denote by $S \circ T$ the family of subsets of G having the form $\{s_i + t_i \mid i \in \{1,\ldots,q\}\}$, where $T' = \{t_1,\ldots,t_q\}$ is any multiset of elements of T having the same cardinality as S. Let $H_1,\ldots,H_r$ be subsets of G such that for any j in $\{1,\ldots,r\}$ we have that $K_j = H_j + \ldots + H_r$ is a subgroup of G and $K_1 = G$. Then, the subgroups K_j, $1 \leqslant j \leqslant r$, yield the following series for G:

$$G = K_1 \supset K_2 \supset \ldots \supset K_r \supset K_{r+1} = 0.$$

Indeed, one has $K_j = H_j + K_{j+1}$, $1 \leqslant j \leqslant r$, $K_r = H_r$. Since $0 \in K_j \cap K_{j+1}$, then K_{j+1} is a subgroup of K_j and H_j is the factor group of K_j by K_{j+1}. Let us consider the

following classes of subsets of G:

$$\Delta = (((0 \circ H_1) + H_2) \circ \ldots + \ldots H_r),$$
$$\Theta = (((0 + H_1) \circ H_2) + \ldots \circ \ldots H_r).$$

Theorem 3.1. *If $T \in \Delta$ and $R \in \Theta$ then (T,R) is a periodic factorization of G. If G is a good group then any factorization (T,R) of G has this form, that is $T \in \Delta$ and $R \in \Theta$.*

Any factorization produced by this method will be called a *Hajós factorization*. One can find a sketch of the proof of Theorem 3.1 in [23]. A complete proof of the second part of this statement can be found in [37]. A variation of this method is presented in [29]. In Section 6, we recall it with a collection of conjectures about the general structure of the factorizations of cyclic groups $\boldsymbol{Z}_n$.

Now, let us consider the definition of Hajós factorizations in the particular case of the cyclic groups. We begin by considering the operation $\circ$ introduced by Hajós, in this specific case. We define an operation in $\boldsymbol{N}\langle a\rangle$, denoted $\circ$, in a consistent way with the relation between subsets of $\boldsymbol{N}$ and polynomials introduced in Section 2.

Definition 3.2. Let $q \in \boldsymbol{N}$ and $T,S = \{s_1,\ldots,s_q\} \subseteq \boldsymbol{N}$. We define a class $a^S \circ a^T$ of polynomials in $\boldsymbol{N}\langle a\rangle$ as follows:

$$a^S \circ a^T = \left\{ \sum_{i=1}^{q} a^{s_i+t_i} \mid T' = \{t_1,\ldots,t_q\} \text{ multiset of elements of } T \right\}.$$

Remark 3.1. Obviously, for all subsets S,T of $\boldsymbol{N}$, we have

$$a^{S \circ T} = a^S \circ a^T.$$

Remark 3.2. Suppose that S,T are subsets of $\boldsymbol{N}$ such that $0 \in T$. Then $S \circ T$ contains S or, equivalently, $a^S \circ a^T$ contains a^S. Indeed, if $0 \in T$, then we can take as T' the multiset of elements of T having the same cardinality as S and all elements equal to 0.

The next lemma gives the form of the subsets H_j which intervene in Hajós' method for cyclic groups.

Lemma 3.2. *Let $H_1,\ldots,H_r$ be subsets of $\boldsymbol{Z}_n$. The following conditions are equivalent:*
(1) *$\forall j \in \{1,\ldots,r\}$, $K_j = H_j + \ldots + H_r$ is a subgroup of $\boldsymbol{Z}_n$ and $K_1 = \boldsymbol{Z}_n$.*
(2) *There exists a chain of divisors of n:*

$$k_0 = 1 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n$$

such that for all $j \in \{1,\ldots,r\}$ we have

$$H_j = \left\{ tk_{j-1} \mid t \in \left\{ 0,\ldots,\frac{k_j}{k_{j-1}} - 1 \right\} \right\}.$$

Proof. Note that condition (1) is verified if and only if one has the following series for $\boldsymbol{Z}_n$:

$$\boldsymbol{Z}_n = K_1 \supset K_2 \supset \ldots \supset K_r \supset K_{r+1} = 0$$

with $K_j = H_j + K_{j+1}$, $1 \leqslant j \leqslant r$, $K_r = H_r$ and H_j is the factor group of K_j by K_{j+1}. Since $\boldsymbol{Z}_n$ is a cyclic group, this is also equivalent to the existence of a chain of divisors of n:

$$k_0 = 1 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n$$

such that, for all $j \in \{1,\ldots,r\}$, we have

$$H_j = \left\{ tk_{j-1} \mid t \in \left\{ 0,\ldots,\frac{k_j}{k_{j-1}} - 1 \right\} \right\}, \quad K_{j+1} = \left\{ tk_j \mid t \in \left\{ 0,\ldots,\frac{n}{k_j} - 1 \right\} \right\}.$$

□

The next proposition describes Hajós' method for cyclic groups.

Proposition 3.2. *Let T,R be subsets of $\boldsymbol{N}$. (T,R) is a Hajós factorization of $\boldsymbol{Z}_n$ if and only if there exists a chain of divisors of n:*

$$k_0 = 1 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n$$

such that

$$a^T \in \left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k_1}-1}{a-1} \right) \cdot \frac{a^{k_2}-1}{a^{k_1}-1} \right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right),$$

$$a^R \in \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_1}-1}{a-1} \right) \circ \frac{a^{k_2}-1}{a^{k_1}-1} \right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right).$$

Proof. By definition, (T,R) is a Hajós factorization of $\boldsymbol{Z}_n$ if and only if there exists $H_1,\ldots,H_r$ subsets of $\boldsymbol{Z}_n$, with $K_j = H_j + \ldots + H_r$ subgroup of $\boldsymbol{Z}_n$, $1 \leqslant j \leqslant r$, $K_1 = \boldsymbol{Z}_n$ and

$$T \in (((0 \circ H_1) + H_2) \circ \ldots + \ldots H_r),$$

$$R \in (((0 + H_1) \circ H_2) + \ldots \circ \ldots H_r).$$

Using Remark 3.1, these two relations hold if and only if we have

$$a^T \in (((1 \circ a^{H_1}) \cdot a^{H_2}) \circ \ldots \cdot \ldots a^{H_r}),$$

$$a^R \in (((1 \cdot a^{H_1}) \circ a^{H_2}) \cdot \ldots \circ \ldots a^{H_r}).$$

Then, using Lemma 3.2, the conclusion holds. □

Now, we can announce one of the main results of this paper which gives a new characterization of Hajós factorizations.

Theorem 3.2. *Let (T,R) be a pair of subsets of $\mathbf{N}$. The following conditions are equivalent*:

(1) *(T,R) is a Hajós factorization of $\mathbf{Z}_n$.*

(2) *There exists a Krasner factorization (I,J) of $\mathbf{Z}_n$ such that (I,R), (T,J) are factorizations of $\mathbf{Z}_n$.*

(3) *There exist $L,M \subseteq \mathbf{N}$ and a Krasner factorization (I,J) of $\mathbf{Z}_n$ such that $a^T = a^I(1 + a^M(a-1))$ and $a^R = a^J(1 + a^L(a-1))$.*

This characterization allows a simpler construction of these factorizations. Indeed, let us consider the following inequalities, where L, M are finite multisets of $\mathbf{N}$ and (I,J) is a Krasner factorization of $\mathbf{Z}_n$:

$$a^I(1 + a^M(a-1)) \geqslant 0, \tag{3.2}$$

$$a^J(1 + a^L(a-1)) \geqslant 0. \tag{3.3}$$

One can prove that there exists a bijection between the solutions L,M of the previous inequalities and the pairs (T,R) of subsets $\mathbf{N}$ verifying condition (3) of Theorem 3.2 (see Proposition A.1). Moreover, there exists a recursive construction of the sets L,M,T,R. Indeed, these inequalities are basic for the construction of a class of finite maximal codes, the so-called d-codes, with $d \leqslant 3$ (see Section 4.3, Propositions 4.2 and 7.1). We will give the complete construction of L,M,T,R in the Appendix. In the following proposition we recall only the part of it which we need in the sequel. It is a direct consequence of Proposition A.1 and Corollary A.1.

Proposition 3.3. *Let L,M,T,R be subsets of $\mathbf{N}$, (I,J) be a Krasner factorization of $\mathbf{Z}_n$. Let k be the smallest divisor greater than 1 in the chain of divisors of n defining (I,J) and (I_1,J_1) the Krasner factorization of $\mathbf{Z}_{n/k}$ such that $J = kI_1 + \{0,\ldots,k-1\}$, $I = kJ_1$. We have*

$$a^T = a^I(1 + a^M(a-1)), \quad a^R = a^J(1 + a^L(a-1))$$

if and only if we have

$$a^T = a^t a^{kR'}, \quad a^R = \sum_{g=0}^{k-1} a^g a^{kT'_g},$$

with $a^{T'_g} = a^{I_1}(1 + a^{M_g}(a-1))$, for all $g \in \{0,\ldots,k-1\}$, $t \in \mathbf{N}$ and $a^{R'} = a^{J_1}(1 + a^{L_1}(a-1))$.

Moreover, let L_1,R_1 be subsets of $\mathbf{N}$ and $t \in \mathbf{N}$. We have

$$a^R = a^t a^{R_1}, \quad a^{R_1} = a^J(1 + a^{L_1}(a-1))$$

if, and only if, we have

$$a^R = a^J(1 + a^L(a-1)), \quad L = \{0,\ldots,t-1\} \cup (L_1 + t).$$

4. Codes

This section is dedicated to the known results about codes and is organized as follows. In Section 4.1 basics on codes are presented. Then, we recall a relationship between maximal codes and factorizations of $\boldsymbol{Z}_n$ proved in [34] (Section 4.2). Finally, Section 4.3 deals with a connection between some factorizing codes and some factorizations of $\boldsymbol{Z}_n$.

4.1. Definitions

In this section we recall some notions about codes which we will use in the sequel (see [2] for a complete survey on this topic and [11] for a list of open problems in this area). A subset C of A^* is a *code* if C is the base of a free submonoid of a free monoid. In other words, C is a code if we have

$$\forall c_1,\dots,c_h,c'_1,\dots,c'_k \in C \quad c_1\cdots c_h = c'_1\cdots c'_k \Rightarrow h = k; \quad \forall i \in \{1,\dots,h\}\ c_i = c'_i.$$

For instance, the set $\{aba, ba, a\}$ is not a code over $A = \{a, b\}$. On the other hand, one can see that the set $\{ba, a\}$ is a code. This is an example of a *prefix* code, C being prefix if $C \cap CA^+ = \emptyset$.

An important class of codes is the class of maximal codes. A code C is *maximal* over A if for any code C' over A we have

$$C \subseteq C' \Rightarrow C = C'.$$

As a basic theorem of Schützenberger shows, a finite code C is maximal if and only if C is *complete*, that is $C^* \cap A^*wA^* \neq \emptyset$, for any $w \in A^*$.

Another related class of codes, introduced by Schützenberger, is the class of the factorizing codes. The definition of such codes is given in terms of polynomials: a code C over A is *factorizing* if there exist two subsets P, S of A^* such that $\underline{C} - 1 = \underline{P}(\underline{A} - 1)\underline{S}$. For instance, a maximal prefix code C is factorizing, by taking $S = \{1\}$ and P equal to the set of the proper prefixes of the words in C. As a special case, note that if C is a factorizing code with P, S *finite*, then C is a finite maximal code; conversely, if C is a finite maximal factorizing code, then P, S are finite sets [2]. However it is not known whether every finite maximal code is factorizing:

Conjecture 4.1. [2, 42] (Schützenberger). *Any finite maximal code is factorizing.*

Some partial results concerning this conjecture are known [5, 15, 18, 22, 33, 35, 43, 47] and weaker forms of it have been posed [31, 32].

4.2. Maximal codes and factorizations of cyclic groups

For a given code C, a *completion* of C is any maximal code C' containing C. If C' is finite, then C' is a *finite completion* of C. By Zorn's lemma, any code C has

a completion but, as it was pointed out by Restivo and Markov [33, 3], there exist finite codes which have no finite completion. The smallest known example is the code $\{a^5, ba^2, ab, b\}$ [2, 33], but the existence of smaller codes having no finite completion has been conjectured [19].

Some relations between maximal codes and factorizations of cyclic groups exist [12, 14, 15, 33, 42]. Here, we are interested in recalling a result in this direction which has been proved in [34]. For this aim, the notion of *pair associated to a code*, introduced in [34], is useful.

Let C be a finite code over $A = \{a, b\}$, such that $b, a^n \in C$, $n \in \boldsymbol{N}$. We associate to C the pair (P, Q) defined as follows:

$$P = \{p \in \boldsymbol{N} \mid b^+a^p \cap C \neq \emptyset\}, \quad Q = \{q \in \boldsymbol{N} \mid a^q b^+ \cap C \neq \emptyset\}.$$

Note that $b \in C$ implies $0 \in P \cap Q$.

For instance, the pair (P, Q) associated to the maximal code $C = \{aa, ab, b\}$ is $P = \{0\}$, $Q = \{0, 1\}$. One can see that $a^2 \in C$ and (P, Q) is a Krasner factorization of $\boldsymbol{Z}_2$. This is a particular case of the result stated in the following proposition.

Proposition 4.1 [34]. *If C is a finite maximal code over $A = \{a, b\}$ such that $b, a^n \in C$, then (P, Q) is a factorization of $\boldsymbol{Z}_n$.*

This result has some consequences for the embedding problem (see Section 6). All known codes having no finite completion are constructed thanks to this relation between maximal codes and factorizations of cyclic groups [17, 26, 33, 34]. We will also use it for our results (see Section 7).

4.3. Factorizing codes and factorizations of cyclic groups

In the previous section we have seen that a factorization of $\boldsymbol{Z}_n$ is canonically associated with some maximal codes. Another result of the same kind is concerned with the so-called d-codes. A *d-code* is a finite maximal code C over $A = \{a, b\}$ such that d is the maximum number of occurrences of b's in the words of C. Moreover there exists a word in C having exactly d occurrences of b's. For instance, $\{ab, aa, b\}$ is a 1-code and $\{a^2, a^2b, ba^{\{0,1\}}b,\ a^3ba, ba^{\{0,1,2\}}ba\}$ is a 2-code.

It was proven that d-codes with $d \leqslant 3$ are factorizing and their structure has been characterized (see [33] for 1-codes, [14, 18] for 2-codes and [15] for 3-codes). They are also studied with respect to the degree and the decomposability [12]. Moreover, it has been proven that for any prime p and for $p = 4$, *(p, p)-codes* (i.e. p-codes C with $b^p \in C$) are factorizing [22].

The structure of 1-, 2- and 3-codes is strictly related to Hajós factorizations of $\boldsymbol{Z}_n$. For the 1-codes, the relation is very simple: Restivo proved that C is a 1-code if and only if $C = a^n + a^l ba^J$, where (I, J) is a Krasner factorization of $\boldsymbol{Z}_n$ [33]. The case of 2- and 3-codes is much more complicated. The construction of these codes is

obtained by the solutions of inequalities (3.2) and (3.3), which, in turn, are related to the Hajós factorizations (see Theorem 3.2). For illustrating this relation, we give here the description of 2-codes. We have reported in Section 7 the statement describing the structure of 3-codes (see Proposition 7.1). We recall that the *reverse* $C^{\sim}$ of C is the set of words of C read from right to left.

Proposition 4.2 [14]. *C is a 2-code if and only if C or $C^{\sim}$ satisfies one of the following equations,*

$$\underline{C} - 1 = a^I(\underline{A} - 1)\left(a^J + \sum_{j\in J} a^{M_j} b a^j\right), \quad \underline{C} - 1 = a^J(\underline{A} - 1)\left(a^I + \sum_{i\in I} a^{L_i} b a^i\right),$$

where (I,J) is a Krasner factorization, M_j is a solution of (3.2), *for any $j \in J$, and L_i is a solution of* (3.3), *for any $i \in I$.*

5. A characterization of Hajós factorizations

In this section, we prove Theorem 3.2 as given in Section 3.3. It gives a simple definition of Hajós factorizations. As a by-product, we obtain that Krasner factorizations can be produced by Hajós' method. We begin by recalling the statement of Theorem 3.2 (see Section 5.1). Then, we prove it in Section 5.3, by using some technical lemmata which are stated in Section 5.2. As we said, by Theorem 3.2, we get a recursive construction of Hajós factorizations for cyclic groups. We give it in Section 5.4 (Proposition 5.1).

5.1. Main result

Let us see the main result concerning Hajós factorizations. In it, we will consider three classes of factorizations of $\boldsymbol{Z}_n$. The first one is the class of factorizations introduced by Hajós. The second one is defined by a property related to Krasner factorizations and the third one is related to 2- and 3-codes. The result below states that all these classes are equal.

Theorem 5.1. *Let (T,R) be a pair of subsets of $\boldsymbol{N}$. The following conditions are equivalent:*

(1) *(T,R) is a Hajós factorization of $\boldsymbol{Z}_n$.*

(2) *There exists a Krasner factorization (I,J) of $\boldsymbol{Z}_n$ such that (I,R), (T,J) are factorizations of $\boldsymbol{Z}_n$.*

(3) *There exist $L,M \subseteq \boldsymbol{N}$ and a Krasner factorization (I,J) of $\boldsymbol{Z}_n$ such that $a^T = a^I(1 + a^M(a-1))$ and $a^R = a^J(1 + a^L(a-1))$.*

5.2. Technical results

The following lemmata will be used for proving Theorem 5.1.

Lemma 5.1. *For any $t \in N$ and $H_2, \ldots, H_r$ subsets of N, $r \in N, r > 2$, we have*

$$a^t((a^{H_2} \circ a^{H_3}) \cdot \ldots a^{H_r}) = (((a^t \cdot a^{H_2}) \circ a^{H_3}) \cdot \ldots a^{H_r}) \tag{5.1}$$

(where in the left and in the right side of this equality, the multiplication and the operation $\circ$ alternate).

Proof. By induction over r. Suppose $r = 3$ and $H_2 = \{h_1, \ldots, h_k\}$. For any multiset $H_3' = \{h_1', \ldots, h_k'\}$ of elements of H_3, we have

$$a^t \sum_{i=1}^{k} a^{h_i + h_i'} = \sum_{i=1}^{k} a^{t + h_i + h_i'}.$$

So (5.1) holds.

Suppose that (5.1) holds for integers less than r, $r > 3$. Let $*$ be either the multiplication or the $\circ$ operation and consider $a^H \in ((a^{H_2} \circ a^{H_3}) \ldots a^{H_{r-1}})$. Since (5.1) holds for $r = 3$ or by the associativity of multiplication, we have $a^t(a^H * a^{H_r}) = (a^t \cdot a^H) * a^{H_r}$. Given that a^H is arbitrary and by the induction hypothesis we have

$$\begin{aligned} a^t \cdot (((a^{H_2} \circ a^{H_3}) \cdot \ldots a^{H_{r-1}}) * a^{H_r}) &= (a^t \cdot ((a^{H_2} \circ a^{H_3}) \cdot \ldots a^{H_{r-1}})) * a^{H_r} \\ &= ((((a^t \cdot a^{H_2}) \circ a^{H_3}) \cdot \ldots a^{H_{r-1}}) * a^{H_r}). \end{aligned}$$

So, (5.1) holds. □

Lemma 5.2. *For any chain of divisors of n,*

$$1 = k_0 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n,$$

we have

$$a^R \in \left(\left(\left(\left(1 \cdot \frac{a^{k_1} - 1}{a - 1}\right) \circ \frac{a^{k_2} - 1}{a^{k_1} - 1}\right) \cdot \frac{a^{k_3} - 1}{a^{k_2} - 1}\right) \circ \ldots \cdot \ldots \frac{a^n - 1}{a^{k_{r-1}} - 1}\right)$$

$$\Leftrightarrow \quad a^R = \sum_{g=0}^{k_1 - 1} a^g a^{R_g} \quad \text{and} \quad \forall g \in \{0, \ldots, k_1 - 1\}$$

$$a^{R_g} \in \left(\left(\left(\frac{a - 1}{a - 1} \circ \frac{a^{k_2} - 1}{a^{k_1} - 1}\right) \cdot \frac{a^{k_3} - 1}{a^{k_2} - 1}\right) \circ \ldots \cdot \ldots \frac{a^n - 1}{a^{k_{r-1}} - 1}\right).$$

Proof. By definition of $\circ$, we have

$$\begin{aligned} &\left(\left(\left(\left(1 \cdot \frac{a^{k_1} - 1}{a - 1}\right) \circ \frac{a^{k_2} - 1}{a^{k_1} - 1}\right) \cdot \frac{a^{k_3} - 1}{a^{k_2} - 1}\right) \circ \ldots \cdot \ldots \frac{a^n - 1}{a^{k_{r-1}} - 1}\right) \\ &= \left(\left(\left(\left(1 \circ \frac{a^{k_2} - 1}{a^{k_1} - 1}\right) + \left(a \circ \frac{a^{k_2} - 1}{a^{k_1} - 1}\right)\right.\right. \right. \\ &\quad \left.\left.\left. + \ldots + \left(a^{k_1 - 1} \circ \frac{a^{k_2} - 1}{a^{k_1} - 1}\right)\right) \cdot \frac{a^{k_3} - 1}{a^{k_2} - 1}\right) \circ \ldots \cdot \ldots \frac{a^n - 1}{a^{k_{r-1}} - 1}\right) \end{aligned}$$

$$= \left(\left(\left(1 \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \frac{a^{k_3}-1}{a^{k_2}-1}\right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right)$$
$$+ \left(\left(\left(a \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \frac{a^{k_3}-1}{a^{k_2}-1}\right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right)$$
$$+ \ldots + \left(\left(\left(a^{k_1-1} \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \frac{a^{k_3}-1}{a^{k_2}-1}\right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right).$$

So, the conclusion follows as we have, using Lemma 5.1,

$$\forall g \in \{0,\ldots,k_1-1\} \quad a^{R'_g} \in \left(\left(\left(a^g \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \frac{a^{k_3}-1}{a^{k_2}-1}\right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right)$$
$$\Leftrightarrow \quad a^{R'_g} = a^g a^{R_g}, \quad a^{R_g} \in \left(\left(\left(1 \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \frac{a^{k_3}-1}{a^{k_2}-1}\right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right).$$

□

5.3. *Proof of the main result*

Now, we prove the main result of this section. In particular, in the proof of the implication (1) ⇒ (2), we show also that any Krasner factorization of $\boldsymbol{Z}_n$ is a Hajós factorization of $\boldsymbol{Z}_n$. Also, one can note that the proof of this implication is a particular application of the method introduced by Obaid (see Section 6, Definition 6.1). Equivalence (2) ⇔ (3) has been proven partially in [14]: here we give a more synthetic proof of it. It states that solutions M, L of inequalities (3.2) and (3.3) are sets of holes of some particular factorizations of $\boldsymbol{Z}_n$.

Proof. (*of Theorem 5.1*). (1) ⇒ (2): According to Proposition 3.2, (T,R) is a Hajós factorization of $\boldsymbol{Z}_n$ if and only if there exists a chain of divisors of n:

$$k_0 = 1 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n$$

such that

$$a^T \in \left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k_1}-1}{a-1}\right) \cdot \frac{a^{k_2}-1}{a^{k_1}-1}\right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right),$$
$$a^R \in \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_1}-1}{a-1}\right) \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right).$$

Let us consider polynomials $P_j = (a^{k_j}-1)/(a^{k_{j-1}}-1)$, for $j \in \{1,\ldots,r\}$. According to Remark 3.2, we have that $a^P \in a^P \circ P_j$, for any subset P of $\boldsymbol{N}$. So, we can choose a^P every time that the operation $\circ$ appears in the expressions:

$$\left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k_1}-1}{a-1}\right) \cdot \frac{a^{k_2}-1}{a^{k_1}-1}\right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right), \tag{*}$$

$$\left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_1}-1}{a-1}\right) \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1}\right). \tag{**}$$

Note that this is equivalent to erasing polynomials P_j, with j odd, from $(*)$ and polynomials P_j, with j even, from $(**)$.

In this way, we will obtain two polynomials a^I and a^J satisfying the following conditions:

$$a^I = \prod_{j \text{ even}, 1 \leqslant j \leqslant r} P_j \in \left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k_1}-1}{a-1} \right) \cdot \frac{a^{k_2}-1}{a^{k_1}-1} \right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right),$$

$$a^J = \prod_{j \text{ odd}, 1 \leqslant j \leqslant r} P_j \in \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_1}-1}{a-1} \right) \circ \frac{a^{k_2}-1}{a^{k_1}-1} \right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right).$$

Then, by definition, (I,J) is a Krasner factorization of $\boldsymbol{Z}_n$ and, by Theorem 3.1, it is a Hajós factorization of $\boldsymbol{Z}_n$. Moreover, by Theorem 3.1 again, (I,R) and (T,J) are factorizations of $\boldsymbol{Z}_n$.

(2) $\Leftrightarrow$ (3): Let (T,R) be a pair of subsets of $\boldsymbol{N}$ and (I,J) a Krasner factorization of $\boldsymbol{Z}_n$. Then, we have

$$a^T = a^I(1 + a^M(a-1)) \geqslant 0, \quad a^R = a^J(1 + a^L(a-1)) \geqslant 0$$
$$\Leftrightarrow \quad a^T a^J = \frac{a^n-1}{a-1}(1 + a^M(a-1)), \quad a^R a^I = \frac{a^n-1}{a-1}(1 + a^L(a-1)).$$

So, using Proposition 3.1, (2) and (3) are equivalent.

(3) $\Rightarrow$ (1): Suppose that there exists $L, M \subsetneq \boldsymbol{N}$ and a Krasner factorization (I,J) of $\boldsymbol{Z}_n$ such that

$$a^T = a^I(1 + a^M(a-1)), \quad a^R = a^J(1 + a^L(a-1)).$$

Note that we can take $T, R \subsetneq \{0, \ldots, n-1\}$ (we proved that (2) and (3) are equivalent).

Let us prove, by induction over the number of the (not necessarily distinct) prime factors of n, that for the chain of divisors of n giving (I,J),

$$1 = k_0 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n,$$

we have

$$a^T \in \left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k_1}-1}{a-1} \right) \cdot \frac{a^{k_2}-1}{a^{k_1}-1} \right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right),$$

$$a^R \in \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_1}-1}{a-1} \right) \circ \frac{a^{k_2}-1}{a^{k_1}-1} \right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right);$$

i.e. (T,R) is a Hajós factorization of $\boldsymbol{Z}_n$ and the first operation in the set containing a^R is a multiplication (see Proposition 3.2).

It is easy to see it for a prime $n = p$. Indeed, in this case, we have $R = \{0, \ldots, p-1\}$ and $T = \{t\}$, for a $t \in \{0, \ldots, p-1\}$. So, we have

$$a^T \in \frac{a-1}{a-1} \circ \frac{a^p-1}{a-1}, \quad a^R \in \frac{a-1}{a-1} \cdot \frac{a^p-1}{a-1},$$

and the conclusion holds.

Let us suppose that the conclusion holds for cyclic groups having the order with a number of prime factors less than a composite number n and let us prove it for n.

If $R = \{0,\dots,n-1\}$ and $T = \{t\}$, for a $t \in \{0,\dots,n-1\}$, we have

$$a^T \in \frac{a-1}{a-1} \circ \frac{a^n-1}{a-1}, \quad a^R \in \frac{a-1}{a-1} \cdot \frac{a^n-1}{a-1},$$

and, again, the conclusion holds.

Otherwise, according to Proposition 3.3 we have

$$J = k_1 I_1 + \{0,\dots,k_1-1\}, \quad I = k_1 J_1, \quad a^T = a^t a^{k_1 s} a^{k_1 R'}, \quad a^R = \sum_{g=0}^{k_1-1} a^g a^{k_1 T'_g},$$

with (I_1,J_1) a Krasner factorization of $\boldsymbol{Z}_{n/k_1}$, $a^{T'_g} = a^{I_1}(1 + a^{M_g}(a-1))$, for all $g \in \{0,\dots,k_1-1\}$, $s \geqslant 0$, $a^s a^{R'} = a^{J_1}(1 + a^{L_1}(a-1))$ and $t \in \{0,\dots,k_1-1\}$.

By using induction hypothesis, $(T'_g, s+R')$ is a Hajós factorization of $\boldsymbol{Z}_{n/k_1}$ and the first operation in the set containing $a^s a^{R'}$ is a multiplication. More precisely, for the chain of divisors of n/k_1 giving (J_1,J_1),

$$k'_1 = 1 \mid k'_2 \mid \dots \mid k'_r = \frac{n}{k_1}, \tag{5.2}$$

we have

$$\forall g \in \{0,\dots,k_1-1\} \quad a^{T'_g} \in \left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k'_2}-1}{a-1}\right) \cdot \frac{a^{k'_3}-1}{a^{k'_2}-1}\right) \circ \dots \cdot \dots \frac{a^{n/k_1}-1}{a^{k'_{r-1}}-1}\right), \tag{5.3}$$

$$a^{s+R'} \in \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k'_2}-1}{a-1}\right) \circ \frac{a^{k'_3}-1}{a^{k'_2}-1}\right) \cdot \dots \circ \dots \frac{a^{n/k_1}-1}{a^{k'_{r-1}}-1}\right). \tag{5.4}$$

Let us consider the chain of divisors of n obtained by setting in (5.2) $k_i = k'_i k_1$, for any $i \in \{1,\dots,r\}$:

$$1 = k_0 \mid k_1 \mid k_2 \mid \dots \mid k_r = n.$$

First, (I,J) is defined by the previous chain. Moreover, with (5.3) and (5.4), we have

$$a^{k_1 T'_g} \in \left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k_2}-1}{a^{k_1}-1}\right) \cdot \frac{a^{k_3}-1}{a^{k_2}-1}\right) \circ \dots \cdot \dots \frac{a^n-1}{a^{k_{r-1}}-1}\right),$$

$$a^{k_1 s} a^{k_1 R'} \in \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_2}-1}{a^{k_1}-1}\right) \circ \frac{a^{k_3}-1}{a^{k_2}-1}\right) \cdot \dots \circ \dots \frac{a^n-1}{a^{k_{r-1}}-1}\right).$$

So, using Lemma 5.1, we have

$$a^T = a^t a^{k_1 s} a^{k_1 R'} \in a^t \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_2}-1}{a^{k_1}-1} \right) \circ \frac{a^{k_3}-1}{a^{k_2}-1} \right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right)$$

$$= \left(\left(\left(a^t \cdot \frac{a^{k_2}-1}{a^{k_1}-1} \right) \circ \frac{a^{k_3}-1}{a^{k_2}-1} \right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right).$$

With this equation, since $a^t \in 1 \circ ((a^{k_1}-1)/(a-1)) = ((a-1)/(a-1)) \circ ((a^{k_1}-1)/(a-1))$, we have

$$a^T \in \left(\left(\left(\frac{a-1}{a-1} \circ \frac{a^{k_1}-1}{a-1} \right) \cdot \frac{a^{k_2}-1}{a^{k_1}-1} \right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right).$$

On the other hand, by using the following equation,

$$\forall g \in \{0, \ldots, k_1 - 1\} \quad a^{k_1 T'_g} \in \left(\left(\left(1 \circ \frac{a^{k_2}-1}{a^{k_1}-1} \right) \cdot \frac{a^{k_3}-1}{a^{k_2}-1} \right) \circ \ldots \cdot \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right),$$

thanks to Lemma 5.2, we have

$$a^R = \sum_{g=0}^{k_1-1} a^g a^{k_1 T'_g} \in \left(\left(\left(\frac{a-1}{a-1} \cdot \frac{a^{k_1}-1}{a-1} \right) \circ \frac{a^{k_2}-1}{a^{k_1}-1} \right) \cdot \ldots \circ \ldots \frac{a^n-1}{a^{k_{r-1}}-1} \right).$$

So, (T,R) is a Hajós factorization of $\boldsymbol{Z}_n$ and the first operation in the set containing a^R is a multiplication. □

Note that, as a by-product of this theorem, we find that solutions M, L of inequalities (3.2) and (3.3) verify another inequality defining the set H of the holes of (T,R):

$$a^H = a^M(a-1)a^L + a^M + a^L \geqslant 0.$$

5.4. A recursive definition of Hajós factorizations

Thanks to Theorem 3.2 and to Proposition 3.3, we can give the following recursive construction of Hajós factorizations.

Proposition 5.1. *The class of Hajós factorizations of $\boldsymbol{Z}_n$ is the smallest class of pairs of subsets of $\boldsymbol{N}$ such that:*

(1) *For any $t, r \in \boldsymbol{N}$, $(\{0\}, \{0\})$ is a Hajós factorization of $\boldsymbol{Z}_1$, with respect to the trivial chain of divisors of* 1.

(2) *(T,R) is a Hajós factorization of $\boldsymbol{Z}_n$, with respect to the chain of divisors of n:*

$$1 = k_0 \mid k_1 \mid k_2 \mid \ldots \mid k_r = n,$$

if, up to translation, we have $T = k_1 R'$, $R = \bigcup_{g=0}^{k_1-1} (k_1 T_g + g)$, where, for any $g \in \{0, \ldots, k_1 - 1\}$, (R', T_g) is a Hajós factorization of $\boldsymbol{Z}_{n/k_1}$, with respect to the chain of divisors of n/k_1:

$$1 = \frac{k_1}{k_1} \mid \frac{k_2}{k_1} \mid \ldots \mid \frac{k_r}{k_1} = \frac{n}{k_1}.$$

6. Some open problems

In this section, we will present some open problems about the structure of the factorizations of an abelian group G. We do this not only to complete the view on the subject, but also to stress the existence of some relations between these open problems and some pre-existing open problems about codes. Indeed, this section is organized in two parts: one concerning factorizations of G, with the particular case of $\boldsymbol{Z}_n$ (Section 6.1), the other concerning codes (Section 6.2). In each case, we prove a partial result about these problems (Propositions 6.1 and 6.4). For this section, a main reference is a paper by Sands [40].

6.1. *Open problems on factorizations*

In his conjecture, Hajós considered the so-called *quasi-periodic* factorizations of G. A factorization (T,R) of G is quasi-periodic if one factor, say T, can be split into disjointed parts T_i, $i \in \{1,\ldots,m\}$, $m > 1$, such that there is a subgroup H of order m and for any $i \in \{1,\ldots,m\}$ there exists $h_i \in H$ with $R + T_i = h_i + (R + T_1)$.

In particular, we have $G = R + T = R + T_1 + H$ and T can be replaced by the periodic factor $T_1 + H$. This fact has two consequences. First, we have the conjecture as to whether it is always possible to replace one factor by a periodic factor. Second, it allows us to stress that any periodic factorization (T,R) is quasi-periodic. In fact, if T is periodic, then $T = T_1 + H$, for a subgroup H of G and a subset T_1 of T (see Section 3.2). In particular, any Hajós factorization is periodic and, so, it is quasi-periodic.

The inclusion of these three classes of factorizations (Hajós, periodic and quasi-periodic) is strict. At the end of this section, we will see an example of a periodic factorization which cannot be obtained by Hajós' method. Moreover, there exist quasi-periodic factorizations which are not periodic, as for instance [21]

$$(T,R) = (\{0,8,16,18,26,34\},\{0,6,12,36,42,48,1,5,25,29,49,53\}).$$

One can see that (T,R) is quasi-periodic, by taking

$$H = \{0,24,48\}, \quad R_1 = \{0,1,5,36\}.$$

Indeed, if we set

$$R_2 = \{6,25,29,42\}, \quad R_3 = \{12,48,49,53\},$$

then we have

$$R_2 + T = R_1 + 24 + T, \quad R_3 + T = R_1 + 48 + T.$$

In particular, $R_1 + H + T = T + R$.

We have the following conjecture:

Conjecture 6.1 [24,21]. *Any factorization of G is quasi-periodic.*

A partial positive result about it was proven by de Bruijn [7]. Moreover Sands [40] proved that this conjecture is false for general abelian groups, but it is still open for cyclic groups [46]. In the same paper [40], Sands stressed certain conditions under which a factorization must be quasi-periodic. For example, this happens if the factor R is contained in a proper subgroup K of G such that G is the direct sum of K and a subgroup H. This remark suggests the following second conjecture about factorizations:

Conjecture 6.2 [40]. *If G is a non-zero finite abelian group and $G = R + T$, where $0 \in R$, $0 \in T$, then R is contained in some proper subgroup K of G.*

For the sake of completeness, let us recall the definition of another class of factorizations, introduced by Obaid [29]. We will conclude this section with a proposition describing the hierarchy of factorizations introduced in this paper, for the particular case of the cyclic group.

Definition 6.1 [29]. Let $H_1, \ldots, H_r$ be subsets of G such that for any i in $\{1, \ldots, r\}$ we have that $K_i = H_i + \ldots + H_r$ is a subgroup of G and $K_1 = G$. Consider the corresponding series for G:

$$G = K_1 \supset K_2 \supset \ldots \supset K_r \supset K_{r+1} = 0.$$

We shall say that the factorization (T, R) arises from the above series if there exists a set $\{h_1, \ldots, h_r \mid h_i \in H_i,\ 1 \leqslant i \leqslant r\}$ such that T (resp. R) is obtained by choosing $C + h_i$, when expression $C \circ H_i$ appears in

$$\Delta = (((0 \circ H_1) + H_2) \circ \ldots + \ldots H_r),$$

$$(\text{resp. } \Theta = (((0 + H_1) \circ H_2) + \ldots \circ \ldots H_r)),$$

i.e. if one has

$$T = H_2 + H_4 + \ldots + h_1 + h_3 + \ldots,$$

$$R = H_1 + H_3 + \ldots + h_2 + h_4 + \ldots \ .$$

In other words, factorizations which arise from the above series can be obtained from Hajós' method if one computes $C \circ D$ by adding a fixed element of D to the set C. However, there are Hajós factorizations which do not arise from the corresponding series of subgroups K_i, $1 \leqslant i \leqslant r$ [29]. We have reported this example in Proposition 6.2.

Let us consider the particular case of the cyclic groups, $\boldsymbol{Z}_n$. As we said, Conjecture 6.1 is still open for this class of groups. Moreover, in this case, Conjecture 6.2 can be formulated as follows:

Conjecture 6.3. *For any factorization (T, R) of $\boldsymbol{Z}_n$ (up to translation) there exists $k \in \boldsymbol{N}$, with $k \neq 1$, $k \mid n$, such that for any $t \in T$ one has $t \equiv 0 \pmod{k}$.*

Another formulation of the previous conjecture arises by considering the following generalization of Hajós' algorithm.

Conjecture 6.4. *For any factorization (T,R) of $\mathbf{Z}_n$ (up to translation), there exists $k \in \mathbf{N}$, with $k \mid n$, $k \neq 1$, and k factorizations $(G,V_0),\ldots,(G,V_{k-1})$ of $\mathbf{Z}_{n/k}$, such that*

$$R = kG, \qquad T = \bigcup_{j=0}^{k-1} (kV_j + j).$$

One can easily notice that Conjectures 6.4 and 6.3 are equivalent [36]. Moreover, using Proposition 5.1, one can see that Hajós factorizations verify Conjecture 6.4.

A first step towards a positive answer to Conjecture 6.3 is given by Proposition 6.1. Indeed, using this result, one can see that if Conjectures 6.3 and 6.1 were equivalent, then they would both be true. This is the main reason for stating the following result.

Proposition 6.1. *If any periodic factorization verifies Conjecture* 6.3, *then any factorization verifies it.*

Proof. Suppose that any periodic factorization verifies Conjecture 6.3. Let us prove that this is the same for any general factorization. The proof is by induction over the number of the (not necessarily) distinct prime factors of n. If n is a prime, then the conclusion holds.

Let us suppose that it holds for cyclic groups of order smaller than $\mathbf{Z}_n$ and let (T,R) be a factorization of $\mathbf{Z}_n$, where n is not prime. If $T = \{0\}$, $R = \{0,\ldots,n-1\}$, then the conclusion holds.

Otherwise, the sum $T + \{0,n,\ldots,(n-1)n\}$ is direct, so by Lemma 3.1, $(T + \{0,n,\ldots,(n-1)n\},R)$ is a periodic factorization of $\mathbf{Z}_{n^2}$. By hypothesis this factorization verifies Conjecture 6.3. Then, there exists $k \mid n^2$, with $k \geqslant 2$, $k \neq n^2$, such that either $T + \{0,n,\ldots,(n-1)n\} = kG$ or $R = kG$. Consequently, there exists $k' \mid n$, with $k' \geqslant 2$, such that either $T + \{0,n,\ldots,(n-1)n\} = k'G'$ or $R = k'G'$.

In the first case there exists $G_1 \subseteq G'$ such that $T = k'G_1$. Thus, in both of the cases, $k' \neq n$ and (T,R) verifies Conjecture 6.3. □

We will conclude this section with a result. In this paper, we have considered several classes of factorizations of $\mathbf{Z}_n$: the class K of Krasner factorizations, the class Ob of factorizations which arise from a series of subgroups of $\mathbf{Z}_n$, H of Hajós factorizations, Pr of periodic factorizations, QP of quasi-periodic factorizations. The next proposition states that each class is strictly included in another one; in particular, Hajós factorizations are not (up to translation) Krasner factorizations.

Proposition 6.2. $K \subset Ob \subset H \subset Pr \subset QP$ (*where all the inclusions are strict*).

Proof. One can easily see that $K \subseteq Ob$: a factorization of $\boldsymbol{Z}_n$ produced by Obaid's method is just a particular translation of a Krasner factorization. We have also observed that $Ob \subseteq H \subseteq Pr \subseteq QP$. So, we have $K \subseteq Ob \subseteq H \subseteq Pr \subseteq QP$. Let us show the strictness of these inclusions. We have already seen a quasi-periodic factorization (T,R) of $\boldsymbol{Z}_{72}$:

$$T = \{0,8,16,18,26,34\}, \quad R = \{0,6,12,36,42,48,1,5,25,29,49,53\}$$

which is not periodic. For getting $K \subset Ob$, it suffices to take $h_i \neq 0$ in an Obaid's factorization. (For instance, consider the factorization $(\{0,2\},\{2,3\})$ of $\boldsymbol{Z}_4$.)

For $Ob \subset H$, one can consider the following factorization (T,R) of the cyclic group $\boldsymbol{Z}_{81}$, reported in [29]:

$$T = \{0,1,2,9,10,11,18,19,47\},$$

$$R = \{0,3,6\} + \{0,27,54\}.$$

Using Proposition 5.1, one has that (T,R) is a Hajós factorization. However, (T,R) cannot be produced by Obaid's method, otherwise it must be a translation of a Krasner factorization which is impossible because of the form of T.

To end the proof of this statement, we just have to show an example of a periodic factorization of $\boldsymbol{Z}_n$ which is not a Hajós factorization.

Let us consider the quasi-periodic factorization (T,R) of $\boldsymbol{Z}_{72}$:

$$T = \{0,8,16,18,26,34\}, \quad R = \{0,6,12,36,42,48,1,5,25,29,49,53\}.$$

If we set $T' = \{0,8,16,18,26,34\} + \{0,1,\ldots,71\} \cdot 72$, then we have that (T',R) is a periodic factorization of $\boldsymbol{Z}_{72^2}$, using Lemma 3.1. One can see that (T',R) is not a Hajós factorization. Indeed, otherwise, using Proposition 5.1 there would exist a divisor k_1 of 72^2, such that either $R = k_1 R'$ or $R = \bigcup_{g=0}^{k_1-1}(k_1 T_g + g)$. The first case cannot hold (5 does not divide 6).

So, suppose that the second case holds. By the same proposition, we have that $k_1 \mid t$, for any $t \in T'$. We must have $k_1 = 2$, so we get $T_0 = \{0,3,6,18,21,24\}$ and $T_1 = \{0,2,12,14,24,26\}$. Again, using Proposition 5.1, $h \neq 1$ must exist, such that $h \mid t$, for any $t \in T_0 \cup T_1$. This is impossible. □

6.2. *Open problems on codes*

We end this section with a conjecture proposed in [34]. This conjecture is concerned with the structure of the factorizations, but it was motivated by the embedding problem for codes. It stresses that one of the reasons for the difficulty of this latter problem might be that the general structure of the factorizations of the cyclic groups is unknown. In [34], the authors use the notion of *unambiguous* pair.

A pair of finite subsets (P,Q) of $\boldsymbol{N}$ is unambiguous (resp. with respect to n) if we have

$$\forall p,p' \in P,\ q,q' \in Q \quad p+q = p'+q' \ (\text{resp.}\,(\text{mod } n)) \Rightarrow p = p', q = q'.$$

Moreover, if (P,Q) is an unambiguous pair and (T,R) is a factorization of $\boldsymbol{Z}_n$, such that $P \subseteq T$ and $Q \subseteq R$, we say that (P,Q) is *embeddable* in (T,R).

For instance, given a finite code $C \subseteq \{a,b\}^*$ such that $a^n, b \in C$, one can associate to it a pair (P,Q) of subsets of $\boldsymbol{N}$ (see Section 4). One can prove that this pair is unambiguous [34]. In addition, if C is maximal, then (P,Q) is a factorization of $\boldsymbol{Z}_n$ (Proposition 4.1). As a corollary, the following result has been proven in [34].

Proposition 6.3. *Let C be a code over $A = \{a,b\}$ such that $b, a^n \in C$. If the associated pair (P,Q) cannot be embedded in a factorization of $\boldsymbol{Z}_n$, then C has no finite completion.*

In [26], Lam gives a unitary framework to the known results in [17, 33, 34]. Indeed, he found a method for constructing some unambiguous pairs of subsets of $\boldsymbol{N}$ which cannot be embedded in any factorization of $\boldsymbol{Z}_n$. So, by using Proposition 6.3, he found a code C containing a^n and having no finite completion for any $n \in \boldsymbol{N} \setminus \{2,3,4\}$. We will give further results in this direction in Section 7.

Conjecture 6.5, proposed in [34], is related to the existence of a special code constructed by Shor [45]. This code, contained in a^*ba^*, arose as a counterexample to a former formulation of factorization conjecture. Moreover, a finite completion of it, if it exists, would be again a counterexample to the latter formulation of the same conjecture. For the pair (P,Q) associated with Shor's code we have $\{0,1\} \subseteq Q$, $\{0,3,8\} \subseteq P$. So, it verifies the hypothesis of the following conjecture:

Conjecture 6.5. *Let $T = \{0,p,k\}$, $R = \{0,1\}$ be an unambiguous pair, where p is a prime which does not divide k. Then (T,R) cannot be embedded in a factorization of a cyclic group.*

If the above conjecture were true, then Shor's code would have no finite completion, using Proposition 6.3. The same holds for the code $a^{\{0,2,5\}}b \cup ba^{\{0,1\}}$, proposed in [34], as the smallest example of code with an associated pair verifying the hypothesis of Conjecture 6.5.

Note also that Conjecture 6.5 is a particular case of Conjecture 6.3. Since Hajós factorizations verify this latter conjecture, then, we have the following partial result about Conjecture 6.5.

Proposition 6.4. *Let $T = \{0,p,k\}$, $R = \{0,1\}$ be an unambiguous pair, where p is a prime which does not divide k. Then (T,R) cannot be embedded in a Hajós factorization. In particular, (T,R) cannot be embedded in a factorization of a good cyclic group.*

Thanks to Proposition 6.3, Proposition 6.4 has some consequences for the embedding problem of codes. We will state this result for codes in Section 7.4.

7. Finite completion for a family of codes

In this section we will prove some results about the embedding problem for finite codes. We begin with the precise definition of the class of codes which we deal with and recall the structure of 3-codes (Section 7.1). Indeed, we need it for our results, with also the characterization of Hajós factorizations given by Theorem 3.2. In Section 7.2, we give a sufficient condition for the existence of a finite completion of our codes and we describe the structure of this completion (see Proposition 7.2). This result has two consequences: the embedding in a factorizing code is decidable for our class of codes (Section 7.3, Proposition 7.3) and the general embedding problem is decidable for a subclass of it (Corollary 7.1). Finally, we end this section with two results: one related to Conjecture 6.5 (Section 7.4, Proposition 7.4) and one related to a question proposed in [26] (Section 7.5, Corollary 7.2).

7.1. Preliminaries

In this section we only consider codes $C \subseteq a^* \cup a^*b \cup ba^*$ with $b \in C$. So, any C has the form $C = a^n \cup a^P b \cup ba^Q$, with $P, Q \subset \boldsymbol{N}$, $n \in \boldsymbol{N}$ and $0 \in P \cap Q$. Our results are proved under the hypothesis that (P, Q) can be embedded in a Hajós factorization of the associated cyclic group $\boldsymbol{Z}_n$. Note that the existence of $n \in \boldsymbol{N}$, for which (P, Q) has the previous property, is decidable. So, Propositions 7.2 and 7.3 can be reformulated for codes with the form $C = a^P b \cup ba^Q$, with $b \in C$. Our results also use the structure of d-codes, $d \leqslant 3$. Proposition 4.2 gives the structure of 2-codes, we recall below the structure of 3-codes.

Proposition 7.1 [15]. *C is a 3-code if and only if C or the reverse of C satisfies one of the following three equations*:

$$\underline{C} - 1 = a^I(\underline{A} - 1)\left(a^J + \sum_{j \in J} a^{M_j} b a^j + \sum_{w \in T} a^{M_w} b w\right)$$

where (I,J) is a Krasner factorization, $T = \sum_{j \in J} a^{M_j} b a^j$ and for any $j \in J$, $w \in T$, M_j, M_w are solutions of (3.2), *or*

$$\underline{C} - 1 = a^J(\underline{A} - 1)\left(a^I + \sum_{i \in I} a^{L_i} b a^i + \sum_{w \in T} a^{L_w} b w\right)$$

where (I,J) is a Krasner factorization, $T = \sum_{i \in I} a^{L_i} b a^i$ and for any $i \in I, w \in T$, L_i, L_w are solutions of (3.3), *or*

$$\underline{C} - 1 = \left(a^J + \sum_{i \in I} a^i b a^{L_i}\right)(\underline{A} - 1)\left(a^J + \sum_{t \in T} a^{M_t} b a^t\right)$$

where (I,J) is a Krasner factorization and for any $i \in I$, L_i is a solution of (3.3). *Moreover, $T \subseteq \bigcup_{i \in I} T_i$, where T_i is defined by $a^{T_i} = a^J(1 + a^{L_i}(a - 1))$. Finally, set*

for $t \in T$, $I_t = \{i \in I \mid t \in T_i\}$, *for any* $i \in I, t \in T$, L_i *and* M_t *are solutions of the following inequalities*: $a^{M_t}(a-1)a^I + a^{L_t} \geqslant 0$, $a^{L_i}(a-1)a^{M_t} + a^{M_t} \geqslant 0$ *if* $t \notin J$, $a^{L_i}(a-1)a^{M_t} + a^{M_t} + a^{L_i} \geqslant 0$ *if* $t \in J$.

7.2. Finite completion

In this section, we give an embedding procedure for codes $C = a^n \cup a^P b \cup ba^Q$, when the pair (P,Q) can be embedded in a Hajós factorization (Proposition 7.2). As a consequence, we prove that if $\boldsymbol{Z}_n$ is a good group, then the embedding problem is decidable for C (Corollary 7.1).

Proposition 7.2. *Let* P, Q *be subsets of* $\boldsymbol{N}$, *with* $0 \in P \cup Q$. *If* (P,Q) *can be embedded in a Hajós factorization* (T,R) *of* $\boldsymbol{Z}_n$, *then the code* $C = a^n \cup a^P b \cup ba^Q$ *can be embedded in the maximal code* C' *defined by*

$$C' = a^n + a^{T\setminus\{0\}}b + a^{I\setminus\{0\}}ba^{J\setminus\{0\}} + ba^R + ba^H b + a^{I\setminus\{0\}}ba^M b + ba^L ba^{J\setminus\{0\}} + ba^L ba^M b.$$

Proof. Suppose that (P,Q) can be embedded in a Hajós factorization (T,R) of $\boldsymbol{Z}n$. Then, according to Theorem 3.2, there exists a Krasner factorization (I,J) of $\boldsymbol{Z}_n$ and a pair of subsets L, M of $\boldsymbol{N}$ such that

$$a^T = a^I(1 + a^M(a-1)), \quad a^R = a^J(1 + a^L(a-1)).$$

By using Proposition 3.1, we also get

$$a^H = a^M(a-1)a^L + a^M + a^L \geqslant 0.$$

Note that $b \in C$ implies $0 \in P \cup Q \subseteq T \cup R$. So, consider C' defined by

$$\begin{aligned} C' &= 1 + (a^I + ba^L)(a + b - 1)(a^J + a^M b) \\ &= a^n + a^{T\setminus\{0\}}b + a^{I\setminus\{0\}}ba^{J\setminus\{0\}} + ba^R + ba^H b + a^{I\setminus\{0\}}ba^M b \\ &\quad + ba^L ba^{J\setminus\{0\}} + ba^L ba^M b. \end{aligned}$$

We have $C \subseteq C'$. Moreover, C' is a maximal code since, using Propositions 4.2 and 7.1, it is a d-code, with $d \leqslant 3$. □

Corollary 7.1. *Let* $\boldsymbol{Z}_n$ *be a good group, let* P, Q *be subsets of* $\boldsymbol{N}$ *and let* $C = a^n \cup a^P b \cup ba^Q$ *be a code containing* b. *The following conditions are equivalent*:

(i) C *has a finite completion.*
(ii) (P,Q) *can be embedded in a factorization of* $\boldsymbol{Z}_n$.
(iii) C *can be embedded in a* d*-code, with* $d \leqslant 3$.

Proof. (i) $\Rightarrow$ (ii) follows from Proposition 6.3. (ii) $\Rightarrow$ (iii) follows from Theorem 3.1 and Proposition 7.2. Finally (iii) $\Rightarrow$ (i) is obvious. □

7.3. *Embedding in a factorizing code*

The next proposition shows that the embedding in a factorizing code is decidable for our codes $C = a^n \cup a^P b \cup ba^Q$. This result gives a relationship between the embedding problem for C and the factorization conjecture, via Hajós factorizations. Indeed, it could be used for constructing possible counterexamples to the factorization conjecture (see Proposition 7.4). On the other hand if the factorization conjecture were true, then the existence of a finite completion for such codes C would be decidable.

Proposition 7.3. *Let P, Q be subsets of $\boldsymbol{N}$, let $n \in \boldsymbol{N}$ and let $C = a^n \cup a^P b \cup ba^Q$ be a code containing b. The following conditions are equivalent:*

(i) *C can be embedded in a finite factorizing code.*

(ii) *(P, Q) can be embedded in a Hajós factorization of $\boldsymbol{Z}_n$.*

(iii) *C can be embedded in a d-code, with $d \leqslant 3$.*

Proof. By using Propositions 4.2, 7.1 and 7.2, we just have to show that (i) $\Rightarrow$ (ii). So, suppose that C can be embedded in a finite factorizing code C'. Then, by definition, there exist finite subsets H, K of A^* such that $\underline{C'} = \underline{H}(\underline{A} - 1)\underline{K} + 1$.

The equation of the words in C' having only one occurrence of b is the following one:

$$a^I b a^J + \sum_{i \in I'} a^i b a^{L_i}(a-1)a^J + \sum_{j \in J'} a^{M_j}(a-1)a^I b a^j,$$

where (I, J) is a Krasner factorization of $\boldsymbol{Z}_n$ and I', J', L_i $(i \in I')$, M_j $(j \in J')$, are subsets of $\boldsymbol{N}$. So, the equation of the words $w \in C' \cap (a^* b \cup ba^*) = a^T b \cup ba^R$ is the following one:

$$a^T b + ba^{R \setminus 0} = ba^{J \setminus 0} + a^{I \setminus 0} b + b + ba^{L_0}(a-1)a^J + a^{M_0}(a-1)a^I b - \sum_{i \in I_1} a^i b - \sum_{j \in J_1} ba^j,$$

where $I_1 = \{i \in I' \setminus 0 \mid (a^{L_i}, 1) > 0\}$, $J_1 = \{j \in J' \setminus 0 \mid (a^{M_j}, 1) > 0\}$.

Moreover, note that we have

$$b \in C \Rightarrow 0 \notin M_0 \cup L_0 \cup I_1 \cup J_1.$$

With this relation, we get

$$0 \leqslant a^T b = a^{I \setminus 0} b + b + a^{M_0}(a-1)a^I b - \sum_{i \in I_1} a^i b \leqslant a^{T_0} b = a^I(1 + a^{M_0}(a-1))b$$

$$\Rightarrow 0 \leqslant a^T \leqslant a^{T_0} = a^I(1 + a^{M_0}(a-1)).$$

Symmetrically, we get

$$0 \leqslant a^{R\setminus 0} \leqslant a^R \leqslant a^{R_0} = a^J(1 + a^{L_0}(a-1)).$$

Consequently, we get

$$P \subseteq T \subseteq T_0, \quad Q \subseteq R \subseteq R_0,$$

where (T_0, R_0) is a Hajós factorization of $\boldsymbol{Z}_n$, using Theorem 3.2. □

Note that (i) $\Rightarrow$ (ii) also holds for codes $C \subseteq a^* \cup a^*ba^*$, containing b, using Proposition 7.3 for $C \cap a^* \cup a^*b \cup ba^*$.

7.4. Examples

In this section, we will give some examples of codes C which do not verify condition (ii) in Proposition 7.3. The main reason for giving these examples is that they enable us to give a partial answer to Conjecture 6.5.

Proposition 7.4. *Let n be a non-negative integer and P, Q subsets of $\boldsymbol{N}$, with $\{0, p, k\} \subseteq P$, $\{0,1\} \subseteq Q$, p being a prime which does not divide k. The code $C = a^n \cup a^P b \cup ba^Q$ has no finite completion which is factorizing. If $\boldsymbol{Z}_n$ is a good group, then C has no finite completion.*

In particular, Shor's code cannot be embedded in a finite maximal code C containing a^n, with $\boldsymbol{Z}_n$ a good group, and it cannot be embedded in a factorizing code.

Remark that Shor's code C does not belong to the family which we deal with, since $C \subseteq a^*ba^*$. However, we get our result, by taking $C \cap (a^*b \cup ba^*)$.

Other codes C which verify Conjecture 6.5 but which cannot be embedded in a factorizing code can be constructed by taking (P, Q) not embeddable in a Hajós factorization (take, as an example, the quasi-periodic factorization of $\boldsymbol{Z}_{72}$ given in Section 6).

7.5. The case $n \in \{2,3,4,6\}$

Finally, the next corollary is related to the following open question posed in [26]: does a finite code having no finite completion and containing a^n, with $n \in \{2,3,4,6\}$, exist? Lam constructed examples of such codes [27]. Corollary 7.2 is a first step for characterizing all codes having this property. Its proof is rather technical and it has been reported in the Appendix.

Corollary 7.2. *Let P, Q be subsets of $\boldsymbol{N}$ and let $C = a^n \cup a^P b \cup ba^Q$ be a code containing b. If $n \in \{2,3,4,6\}$, then C can be embedded in a d-code, with $d \leqslant 3$.*

Appendix A

In this appendix we have gathered two technical proofs. The first one concerns solutions of inequalities (3.2) and (3.3):

$$a^{(M,I)} = a^I(1 + a^M(a-1)), \tag{3.2}$$

$$a^{(L,J)} = a^J(1 + a^L(a-1)), \tag{3.3}$$

where L, M are finite multisets of $\boldsymbol{N}$ and (I,J) is a Krasner factorization. First, in Proposition A.1, we recall the characterization of their solutions, as it was given in [14]. Then, in Corollary A.1, we prove a different formulation of one condition in Proposition A.1. Finally, we prove Corollary 7.2.

Proposition A.1 [14]. (i) *M is a solution of the inequation $a^M \leqslant a^{M+1} + 1$ if and only if either $M = \emptyset$ or $M = \{0,\ldots,t\}$, with $t \in \boldsymbol{N}$. In the second case we have $a^{M+1} + 1 - a^M = a^{t+1}$.*

(ii) *M is a solution of* (3.2) (*resp.* (3.3)) *having min $M = 0$ if and only if we have $M = \{0,\ldots,t\} \cup (M' + t + 1)$, with $t \in \boldsymbol{N}$, M' a solution of* (3.2) (*resp.* (3.3)) *and either $M' = \emptyset$ or min $M' > 0$. Moreover we have*

$$a^{(M,I)} = a^{t+1}a^{(M',I)} \quad (\textit{resp. } a^{(M,J)} = a^{t+1}a^{(M',J)}).$$

(iii) *M is a solution of* (3.2) *having min $M > 0$ if and only if $M = kL + \{0,\ldots,k-1\}$, where L is a solution of* (3.3) *when we substitute J with J_1. Moreover, for any solution L of the former inequation, M given by the previous equation is a solution of* (3.2) *and we have*

$$a^{(M,I)} = a^{k(L,J_1)}.$$

(iv) *L is a solution of* (3.3) *if and only if we have $L = \bigcup_{g=0}^{k-1}(kM_g + g)$, where M_g is a solution of* (3.2) *when we substitute I with I_1. Moreover we have*

$$a^{(L,J)} = a^{\cup_{g=0}^{k-1}(k(M_g,I_1)+g)}.$$

Corollary A.1 *Let L_1, R_1 be subsets of $\boldsymbol{N}$ and $t \in \boldsymbol{N}$. We have*

$$a^R = a^t a^{R_1}, \quad a^{R_1} = a^J(1 + a^{L_1}(a-1))$$

if, and only if, we have

$$a^R = a^J(1 + a^L(a-1)), \quad L = \{0,\ldots,t-1\} \cup (L_1 + t).$$

Proof. Suppose that we have

$$a^R = a^t a^{R_1} = a^t a^J(1 + a^{L_1}(a-1)).$$

Using this relation, we have

$$a^R = a^J(a^t + a^t a^{L_1}(a-1)) = a^J(1 + a^{\{0,\dots,t-1\}}(a-1) + a^t a^{L_1}(a-1)),$$

which implies

$$a^R = a^J(1 + a^L(a-1)), \quad L = \{0,\dots,t-1\} \cup (L_1 + t).$$

Conversely, suppose that we have

$$a^R = a^J(1 + a^L(a-1)), \quad L = \{0,\dots,t-1\} \cup (L_1 + t).$$

Then, using Proposition A.1(ii), we have

$$L = \{0,\dots,t'-1\} \cup (L_1' + t')$$

with $t' \in N$, $t' \geqslant t$, *min* $L_1' > 0$, L_1' a solution of (3.3). Consequently, we have

$$L_1 = \{0,\dots,t'-t-1\} \cup (L_1' + t' - t).$$

Moreover, using Proposition A.1(ii), L_1 is a solution of (3.3) and we have

$$a^R = a^t a^{R_1}, \quad a^{R_1} = a^J(1 + a^{L_1}(a-1)). \qquad \square$$

Proof. (*of Corollary 7.2*). By our hypotheses, (P,Q) is an unambiguous pair (with respect to n). We will prove that if $n \in \{2,3,4,6\}$, then (P,Q) can be embedded in a factorization of $\boldsymbol{Z}_n$. Then, by Corollary 7.1, C can be embedded in a d-code with $d \leqslant 3$.

We need the following remarks:

(i) If $|P| \cdot |Q| > n$, then (P,Q) cannot be unambiguous (with respect to n).

(ii) If $|P| \cdot |Q| = n$ and (P,Q) is an unambiguous pair (with respect to n), then (P,Q) is a factorization of $\boldsymbol{Z}_n$.

(iii) We have $0 \in P \cap Q$ and we can take $P \cup Q \subseteq \{0,\dots,n-1\}$.

(iv) An unambiguous pair (with respect to n) $(P,\{0\})$ (resp. $(\{0\},Q)$) of $\boldsymbol{Z}_n$ can be embedded in the factorization $(\{0,\dots,n-1\},\{0\})$ (resp. $(\{0\},\{0,\dots,n-1\})$) of $\boldsymbol{Z}_n$.

Case $n = 2$: By using (i) we have $|P| \leqslant 2$ and $|Q| = 1$ or $|P| = 1$ and $|Q| \leqslant 2$. In both the cases, by (iv), (P,Q) can be embedded in a factorization of $\boldsymbol{Z}_2$.

Case $n = 3$: By using (i) we have $|Q| = 1$, $|P| \leqslant 3$ or $|P| = 1$ and $|Q| \leqslant 3$. In both the cases, by (iv), (P,Q) can be embedded in a factorization of $\boldsymbol{Z}_3$.

Case $n = 4$: By using (i) we have two cases:

(1) $|P| = 1$, $|Q| \leqslant 4$ or $|Q| = 1$, $|P| \leqslant 4$.

(2) $|P| = |Q| = 2$.

By using (iv) in the first case and by using (ii) in the second case, (P,Q) can be embedded in a factorization of $\boldsymbol{Z}_4$.

Case $n = 6$. By using (i) we have three cases:

(1) $|P| = 1$, $|Q| \leqslant 6$ or $|Q| = 1$, $|P| \leqslant 6$.

(2) $|P| = 3$, $|Q| = 2$ or $|Q| = 3$, $|P| = 2$.

(3) $|P| = |Q| = 2$.

By using (iv) in the first case and by using (ii) in the second case, (P, Q) can be embedded in a factorization of Z_6.

Finally one can see that the pairs (P, Q) which verify case (3) are the following:

$(\{0,1\},\{0,2\})$, $(\{0,1\},\{0,3\})$, $(\{0,1\},\{0,4\})$, $(\{0,2\},\{0,3\})$,

$(\{0,2\},\{0,5\})$, $(\{0,3\},\{0,4\})$, $(\{0,3\},\{0,5\})$, $(\{0,4\},\{0,5\})$.

Each of them can be embedded in one of the following factorizations of Z_6:

$(\{0,1\},\{0,2,4\})$, $(\{0,1,2\},\{0,3\})$, $(\{0,2,4\},\{0,5\})$, $(\{0,3\},\{0,4,5\})$. □

Acknowledgements

It is a pleasure for me to thank V. Bruyère and D. Perrin for useful suggestions and for pointing out some of the references of factorizations. Thanks are also due to A. De Santis and in particular to the anonymous referee for improving the presentation of the paper. Thanks to A. de Luca, A. Restivo and C. Reutenauer for their encouragement.

References

[1] J. Ashley, B. Marcus, D. Perrin and S. Tuncel, Surjective extensions of sliding block-codes, LITP Report 91–23, 1991.

[2] J. Berstel and D. Perrin, *Theory of Codes* (Academic Press, New York, 1985).

[3] J. Berstel and D. Perrin, Trends in the theory of codes, *Bull. EATCS* **29** (1986) 84–95.

[4] J. Berstel and C. Reutenauer, *Rational Series and their Languages*, EATCS Monographs, Vol. 12 (Springer, Berlin, 1988).

[5] J.M. Boë, Factorisation par excés du monoïde libre, LIRMM Report 94-005, 1994.

[6] N.G. de Bruijn, On the factorization of finite abelian groups, *Indag. Math. Kon. Ned. Akad. Wetensch. Amsterdam* **15** (1953) 258–264.

[7] N.G. de Bruijn, On the factorization of cyclic groups, *Indag. Math. Kon. Ned. Akad. Wetensch. Amsterdam* **15** (1953) 370–377.

[8] V. Bruyère, Codes, Thesis, Université de Mons-Hainaut, 1991.

[9] V. Bruyère, Completion of codes, in: M. Ito, ed., *Proc. Coll. on Words, Languages and Combinatorics* (World Scientific, Singapore, 1992) 30–34.

[10] V. Bruyère, Automata and codes with bounded deciphering delay, *Proc. Latin'92*, Lecture Notes in Computer Science, Vol. 583 (Springer, Berlin, 1992) 99–107.

[11] V. Bruyère, Research topics in the theory of codes, *Bull. EATCS* **48** (1992) 412–424.

[12] V. Bruyère and C. De Felice, Degree and decomposability of variable-length codes, *Proc. ICALP'91*, Lecture Notes in Computer Science, Vol. 510 (Springer, Berlin, 1991) 575–587.

[13] V. Bruyère, L.M. Wang and L. Zhang, On completion of codes with finite deciphering delay, *European J. Combin.* **11** (1990) 513–521.

[14] C. De Felice, Construction of a family of finite maximal codes, *Theoret. Comput. Sci.* **63** (1989) 157–184; *Proc. STACS'88*, Lecture Notes in Computer Science, Vol. 294 (Springer, Berlin, 1988) 159–169.
[15] C. De Felice, A partial result about the factorization conjecture for finite variable-length codes, *Discrete Math.* **122** (1993) 137–152.
[16] C. De Felice, Completing codes by Hajós factorizations of groups, *Proc. Conf. on Semigroups, Automata and Languages*, Portugal, 1994 (World Scientific, Singapore, to appear).
[17] C. De Felice and A. Restivo, Some results on finite maximal codes, *RAIRO Inform. Théor.* **19** (1985) 383–403.
[18] C. De Felice and C. Reutenauer, Solution partielle de la conjecture de factorisation des codes, *C.R. Acad. Sci. Paris* **302** (1986) 169–170.
[19] D. Derencourt, A three-word code which is not prefix–suffix composed, Rapport LIFL IT 265, 1995.
[20] A. Ehrenfeucht and G. Rozenberg, Each regular code is included in a regular maximal code, *RAIRO Inform. Théor.* **20** (1985) 89–96.
[21] L. Fuchs, *Abelian Groups* (Pergamon Press, Oxford, 1960).
[22] C. Gu and L. Zhang, Two classes of factorizing codes – (p, p)-codes and $(4,4)$-codes, in: M. Ito, ed., *Proc. Coll. on Words, Languages and Combinatorics* (World Scientific, Singapore, 1992).
[23] G. Hajós, Sur la factorisation des groupes abéliens, *Casopis Pest. Mat. Fys.* **74** (1950) 157–162.
[24] G. Hajós, Sur le problème de factorisation des groupes cycliques, *Acta Math. Acad. Sc. Hungaricae* **1** (1950) 189–195.
[25] M. Krasner and B. Ranulac, Sur une propriété des polynômes de la division du cercle, *C.R. Acad. Sci. Paris* **240** (1937) 397–399.
[26] N.H. Lam, On codes having no finite completion, *Proc. STACS'94*, Lecture Notes in Computer Science, Vol. 775 (Springer, Berlin, 1994) 691–698; *RAIRO Inform. Théor.* **29** (1995) 145–155.
[27] N.H. Lam, A note on codes having no finite completions, *Inf. Proc. Lett.* **55** (1995) 185–188.
[28] R. Montalbano, Local automata and completion, *Proc. STACS'93*, Lecture Notes in Computer Science, Vol. 665 (Springer, Berlin, 1993) 333–342.
[29] E.E. Obaid, On a variation of Sands' method, *Internat. J. Math. Math. Sci.* **9** (1986) 597–604.
[30] D. Perrin, Completing biprefix codes, *Theoret. Comput. Sci.* **28** (1984) 329–336.
[31] D. Perrin and M.P. Schützenberger, Un problème élémentaire de la théorie de l'information, *Théorie de l'Information Coll. Internat. CNRS*, Vol. 276, Cachan (1977) 249–260.
[32] D. Perrin and M.P. Schützenberger, A conjecture on sets of differences of integer pairs, *J. Combin. Theory Ser. B* **30** (1981) 91–93.
[33] A. Restivo, On codes having no finite completions, *Discrete Math.* **17** (1977) 309–316.
[34] A. Restivo, S. Salemi and T. Sportelli, Completing codes, *RAIRO Inform. Théor.* **23** (1989) 135–147.
[35] C. Reutenauer, Non commutative factorization of variable-length codes, *J. Pure Appl. Algebra* **36** (1985) 167–186.
[36] C. Reutenauer, private communication.
[37] A.D. Sands, On the factorisation of finite abelian groups, *Acta Math. Acad. Sc. Hungaricae* **8** (1957) 65–86.
[38] A.D. Sands, The factorization of abelian groups, *Quart. J. Math. Oxford Series 2* **10** (1959) 81–91.
[39] A.D. Sands, Factorization of cyclic groups, *Proc. Coll. on Abelian Groups* (Akad. Kiadó Budapest, 1964).
[40] A.D. Sands, On a conjecture of G. Hajós, *Glasgow Math. J.* **15** (1974) 88–89.
[41] M.P. Schützenberger, Une théorie algébrique du codage, *Séminaire Dubreil-Pisot* (1955–56) Exposé no. 15.
[42] M.P. Schützenberger, Codes à longueur variable, manuscript, 1965; reprinted in: D. Perrin, ed., *Théorie des Codes*, LITP (1979) 247–271.
[43] M.P. Schützenberger, Sur certains sous-monoïdes libres, *Bull. Soc. Math. France* **93** (1965) 209–223.
[44] Z. Shen and L. Zhang, Completion of recognizable bifix codes, *Theoret. Comput. Sci.*, to appear.
[45] P. Shor, A counterexample to the triangle conjecture, *J. Combin. Theory Ser. A* **38** (1985) 110–112.
[46] S. Szabo, private communication.
[47] L. Zhang, private communication.